DOCUMENT RESUME

ED 432 686 CE 079 041

AUTHOR	Athanasou, James A.
TITLE	Analysis of Responses to Vocational Interest Items: A Study of Australian High School Students. Occasional Paper Number 10.
INSTITUTION	Technology Univ., Sydney (Australia).
PUB DATE	1999-08-00
NOTE	20p.
PUB TYPE	Reports - Research (143)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	Educational Research; Foreign Countries; *High School Students; High Schools; *Interest Inventories; Occupational Tests; Predictive Measurement; *Scores; *Test Validity; Vocational Education; *Vocational Interests; *Work Environment
IDENTIFIERS	*Australia; Hollands Theory of Occupational Choice

ABSTRACT

A study used item response theory to examine the vocational interests of 2,709 high school students (1,436 males; 1,273 females) in Australia in relation to Holland's vocational interest typology (1973, 1985, 1997), which identifies six fundamental vocational types (realistic, investigative, artistic, social, enterprising, and conventional) that link interests and work environments within the ambit of personality. Participants completed a 24-item questionnaire. Partial credit analysis was used to determine the location of the four questions that made up each of the vocational scales. Infit mean squares centered on 1.0 and separability was satisfactory for all scales (0.85 to 0.99) except the investigative scale. The conclusion was that scales and items conformed generally to the measurement model. The analysis of items using a Rasch model provided new information on how individuals responded to items and the complexity of responses within interest categories. It was concluded that simple raw scores or summing scores may not offer a valid basis for assessment of interests. (Included are two tables and six figures that show the distances between rating scales on the six scales. Contains 15 references.) (YLB)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED 432 686

OCCASIONAL PAPER
Number 10

ANALYSIS OF RESPONSES TO VOCATIONAL INTEREST ITEMS: A STUDY OF AUSTRALIAN HIGH SCHOOL STUDENTS

James A Athanasou
Faculty of Education
University of Technology, Sydney

U.S. DEPARTMENT OF EDUCATION
Office of Educational Research and Improvement
EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)
This document has been reproduced as received from the person or organization originating it.
Minor changes have been made to improve reproduction quality.
Points of view or opinions stated in this document do not necessarily represent official OERI position or policy.

August 1999

BEST COPY AVAILABLE

CE079041

ABSTRACT

The vocational interests of 2709 male and female high school students in Australia were examined using item response theory. The present study used a national probability sample of Australian youth. Participants completed a 24-item questionnaire that reflected the vocational interest typology of Holland. Partial credit analysis was used to determine the location of the 4 questions that comprised each of the vocational scales. Infit mean squares centred on 1.0 and separability was satisfactory for all scales (0.85 to 0.99) except the Investigative scale. It was concluded that scales and items conformed generally to the measurement model. The analysis of items using a Rasch model provided new information on how individuals responded to items and the complexity of responses within interest categories. It is argued that simple raw scores or summing scores may not offer a valid basis for assessment of interests.

BEST COPY AVAILABLE

ANALYSIS OF RESPONSES TO VOCATIONAL INTEREST ITEMS: A STUDY OF AUSTRALIAN HIGH SCHOOL STUDENTS

The role of vocational interests in learning and career development has been supported over some 80 years through the pioneering work of Thurstone, Strong, Kuder, Roe, Holland and others. Interest is a robust construct that has been linked with educational choices, vocational development, workplace performance, job satisfaction, and personality characteristics. Nevertheless, career behaviours have sometimes shown no correlations or weak correlations with interests and challenged any hypothesised congruence or expected links. Such correlational studies have typically relied on interest scales with high internal consistency containing homogeneous items that reflect a vocational interest dimension.

The traditional interest scales might show weak correlations when the items include some in which there are inconsistencies amongst people as to where these items fit on an interest dimension. Furthermore, since every interest scale represents a particular selection of potential interest items administered to a trial group of participants then the values are unstable because of the particular characteristics of the group(s) and the score estimates are restricted by the particular characteristics of the resulting items chosen. It has always been assumed that the extent of vocational interest in an area can be scaled easily along a dimension based on raw scores simply by adding numbers. Yet the psychometric properties of these numbers and the validity of items that make up interest categories have not been investigated. Consistent with classical measurement theory, the development of many interest scales has been based to a great extent on item-total correlations to produce homogeneous and internally consistent dimensions (exceptions to this tradition are the empirically keyed occupational scales of interests).

Scales based on a Rasch measurement model offer an alternative and Embertson (1996) has pointed out how these newer measurement models have become "mainstream as a theoretical basis for psychological measurement" (p. 341). While item response theory has found wide application in the assessment of ability, popular achievement tests, attitude scales and more recently even to personality measures, it has not been implemented in the field of career interests.

There are a number of advantages in using Rasch estimates for interest items. These include: (a) interest items can be located on an interval scale; and (b) the person's level of interest can represented on the same dimension as the items. Consequently it is possible to determine whether the persons responding to an interest questionnaire are matched to the items on the interest scale. Furthermore for developmental studies, changes in interest can be mapped on a scale. Finally it is possible to predict the chances of a person being interested in items (such as occupations, courses, activities) in addition to those on which he or she has been assessed.

One reason for the lack of application of item-response theory might be the general satisfaction of vocational researchers with existing instruments that have shown their robustness in guidance and counselling over many years. Secondly, interest questionnaires have a considerable pedigree of application and theory in the practice of career assessment and there may well be a reluctance to harness complex probabilistic measurement models to already quite popular scales. Thirdly, many interest items are typically scored on a rating scale rather than on a Yes/No basis and the application of Rasch models to polytomously-scored items although some 30 years old, has been a relatively recent development in applied psychometrics.

An early study by Elton and Rose (1975) applied Rasch scaling to the *Vocational Preference Inventory*, in which the items are scored Yes/No. This was undertaken in order to produce a sex-free form of the inventory. A search of the literature revealed that this was the only application of a Rasch model to interests. Yet, in other contexts guidelines for the development of scales that are intended to assess a construct have been well-established (see Waugh 1999, p.67; Wright and Masters, 1981).

The focus of the present study is on the nature of the item responses that constitute interest scales. The purpose is to explore the application of a Rasch measurement model to polytomously-scored items on interest scales based on the typology of Holland (1996) and used with an Australian cohort of high school students. The following sections describe some contexts for this study.

Holland's vocational typology

Holland (1973, 1985, 1997) has identified six fundamental vocational types (Realistic, Investigative, Artistic, Social, Enterprising and Conventional) that link interests and work environments within the ambit of personality. Combinations of types and their

interaction with environments form the foundation for a comprehensive model of vocational choice with significant predictions for satisfaction and adjustment. This approach is now one of the most widely cited theories of career development, with considerable application to careers guidance and counselling (Borgen, 1991) and it has been a major influence on vocational research in Australia (see Lokan & Taylor, 1986). The reader is referred to the latest exposition of the theory (Holland, 1997; see also Reardon & Lenz, 1998,Chapters 2-3).

The six types were assessed originally by the *Vocational Preference Inventory* and more recently by the *Self-Directed Search* (adapted for use in Australia by the Australian Council for Educational Research) as well as being applied to other vocational assessments such as the *Strong Interest Inventory*. Related measures have been developed for research purposes, such as in Australian studies of subject-choice (Ainley, Robinson, Harvey-Beavis, Elsworth & Fleming, 1994). All of these measures have relied on the summation of raw scores or ratings to form a scale score for each interest category.

Youth in Transition

Youth in Transition is part of the Longitudinal Surveys of Australian Youth conducted for the Federal Government and it seeks to map the vocational, educational and social pathways of young Australians from high school and beyond. The surveys are made up of four cohorts of young people born in 1961, 1965, 1970 and 1975. They involve a two-stage stratified probability sample of 25 students from a nationwide sample of government, independent and Catholic school systems. The 1970 cohort was used in this study and at the outset comprised 5,473 10-year-olds who were first assessed in 1980, and then followed up at yearly intervals from 1985-1994 (further details are provided in the Methods section). In their review of longitudinal studies, Lamb, Polesel and Teese (1995, p.27) indicated that "...it represents one of the most substantial long-term studies of outcomes undertaken in Australia".

Research issues

In this study, the key research issue was to describe how well the preferences of this sample of high school students on a set of interest items were represented in the six Holland scales. These were analysed in terms of item-response theory, that is, as scale-free measures and with sample-free item difficulties. For instance, it is possible

BEST COPY AVAILABLE

to determine the extent to which increasing levels of overall scientific interest are required in order to respond to different categories on a 4-point scale, that is, from a rating of 1 for 'dislike very much' to a rating of 4 for 'like very much' for items such as "doing all kinds of experiments". The variation between observed and expected response patterns can be used to indicate the compatibility of the questionnaire data and a hypothetical item-response model for each of the six Holland types. Support for the validity of an interest scale would depend, *inter alia*, upon the extent to which errors are low and students are spread out along an interest dimension; the fit with the measurement model; whether the amount of interest required to pass from one scale category to the next (ie., from 'like somewhat' to 'like very much') is ordered; and any theoretical ideas supporting the interest category and the items on the scale.

METHOD

Participants. The participants in this study comprised 2,709 students (males=1436; female=1273) from the 1970 Youth in Transition study cohort, who were first tested as part of the Australian Studies of School Performance in 1980. When contacted again in 1985 for the first time, some 2,709 out of 3,294 responded completely to every item in the interest questionnaire and were included in this study. The mean age of the sample was 15.5 months (SD=0.3).

Instrument. The interest inventory used in this study was a 24-item questionnaire of the Holland typology of interests developed especially for administration by mail. It formed one of the twelve sections of the larger survey. Students were asked 'How do you feel about each of these activities?' and responded on a four point scale from 'like very much' to 'like somewhat' through to 'dislike somewhat' and 'dislike very much' for items such as: working with machines and tools (R), doing all kinds of experiments (I), acting in plays (A), helping others (S), managing other people (E) and doing office work (C). (see Australian Council for Educational Research, Longitudinal Surveys of Australian Youth, Technical Paper Number 5 for a complete copy of the survey questionnaire). Due to restrictions of both space and response time the questionnaire was limited to four items per scale and designed for moderate levels of internal consistency with alpha coefficients for the six RIASEC scales of 0.802, 0.602, 0.636, 0.545, 0.641, and 0.704 respectively. The questionnaire has been used

BEST COPY AVAILABLE

subsequently in other large-scale studies and validated against subject choice (Ainley et al., 1994).

Analysis. Partial credit analysis (Wright & Masters, 1982) of each of the six scales and their items was undertaken with *Quest* (Adams & Khoo, 1994). The resultant logit values represent an interval scale of the log odds of students agreeing with an item from those easiest (negative logit values) to those hardest with which to agree (positive logit values). Threshold values are calculated to indicate the probability of passing from one rating to the next (e.g., from 'dislike' to 'like somewhat' or from 'like somewhat' to 'like very much'). The fit of the responses to the measurement model is determined on the basis of infit and outfit statistics. These have an expected value of 1 and usually range from 0.75 to 1.3. Reliability was calculated by a Separability Index (with a value of 1 representing high separability). For the purposes of the Quest program the ratings 'dislike somewhat' and 'dislike very much' were combined into one group. Further details of the analysis are described in the relevant sections of the results.

RESULTS

The results are set out in Figures 1(a) – 1(f) and Tables 1-2. Table 1 includes the statistics for the six interest scales and Table 2 includes the items and their difficulties on the inventory scale. Figure 1 is an item-ability map that sets out the student interests and the item difficulties on the same calibrated scale with zero representing the mean of the item difficulties.

Insert Tables 1-2 about here

Interest scales

Table 1 lists the basic psychometric statistics relating to the six scales. Firstly, the variation within each scale indicates considerable differences (the standard deviations for the scales varied from 0.08 to 1.12). Secondly, the separability reliability index is adequate for all scales except the Investigative scale. Thirdly, examination of the infit and outfit mean squares is generally consistent with the model. The expected value of mean squares is 1: with this group the infit mean squares ranged from 0.98 to 1.10.

Outfit mean squares were also acceptable except for the outfit mean square of 1.35 for the Realistic scale, however outfit statistics include every response even outliers or extreme observations.

Items

The threshold values for each item and rating are listed in Table 2. These values are consistent with the measurement model as the ratings for each item represent an ordered category of responses.

Item-ability maps

It may be helpful to take the four item Realistic scale as an example for interpretation of the item-ability maps. Each X in Figure 1(a) represents 20 students and the items 1.1 (dislike somewhat very much), 1.2 (like somewhat) and 1.3 (like very much) represent the scaled responses to the first question. The placement of students and items on the same scale allows one to consider how well the four different items and each of their ratings (dislike to like very much) matched the students' range of interests.

Insert Figure 1 about here

Students' interests ranged from around –2 to +3 logits and the difficulties ranged from around –3 (lowest realistic interest – dislikes driving cars) to +3 logits (highest realistic interest – likes repairing things very much). The positive logit values represent the items that demand the highest levels of Realistic interest. Note that a dislike of driving (2.1) and liking driving somewhat (2.2) are well below the level of most students' Realistic interests. One would need an extremely high interest in driving to account even for a moderate level of Realistic interest. Indeed, most items were generally below the level of Realistic interest of the group. Each of the subsequent scales can be interpreted in a similar manner.

DISCUSSION AND CONCLUSIONS

The analysis of these responses using a Rasch model (ie., partial credit analysis) provided an alternative means of describing and calibrating students' interest responses on the six Holland dimensions. The results showed intricate relationships between a person's overall level of a vocational interest and the probability of his/her endorsing a rating category ascribed to particular items.

At the level of the six scales there was a reasonable fit to the measurement model. For instance the infit and outfit mean squares centred on 1.0 and were generally within the range of 0.75 to 1.3. The thresholds for each item were also consistent with the measurement model and values were ordered from 'dislike' to 'like very much'. Broadly similar comments can be made about the 24 items with some exceptions.

However, there is evidence that ratings are not unequivocal indicators of interest. The ratings for each item tap different levels of interest and the distances between any two of the rating categories vary considerably across these 24 items. For instance, the distance between 'like somewhat' and 'dislike' on the Realistic scale alone varies from 1.68 (item 1) through 0.88 (item 2) to 2.14 (item 3) and 2.05 (item 4). Although Likert scales assume that raw scores can be added to produce a quantitative index it is not clear that similar quantities are being added within each of the scales. Across the six RIASEC scales there are additional problems in that the items and their ratings are not always matched with the ability (ie., interest) of the sample. A clear example of this is seen in a comparison of the ability-item maps of the Artistic and Social scales (Figures 1c and 1d).

On the Realistic scale there is scope for some items that tap higher levels of interest and this is seen in the item-ability map (Figure 1a). The Investigative scale suffers from low separability and reference to Figure 1(b) indicates that the cause lies in the fact that the average difficulty for three groups of 2-3 rating categories is identical. Artistic interests are represented by a large number of people responding similarly to items with ratings of 'like very much' and Figure 1(c) supports the need for ratings that would tap some higher levels of interest. The interest items on the Social scale (Figure 1d) suffer from serious ceiling effects with many items below the 'ability' level of the group. Amongst the six scales, the Enterprising scale (Figure 1e)

is distinguished by a better targeting of items whereas the Conventional scale (Figure 1f) shows evidence of students who overlap the items at both extremes.

The 24 items for these scales provided a unique and meaningful context for the analysis of responses to interest items in a questionnaire. It may be helpful again to emphasise that it was not the purpose of this report to comment on these scales *per se* but merely to use them as an example of the application of Rasch analysis in order to describe the intricate patterns of item responses that can affect scale scores on vocational interest dimensions. Their construction reflected a classical measurement model with an emphasis on raw scores as the basis for the formation of scales yet there is evidence that even in such a carefully constructed questionnaire the addition of interest ratings may not be justified. Fortunately, in this study raw scores within a scale would correlate highly with a scale score based on logits but some inter-scale comparisons might be fraught with problems. The same raw score can represent vastly different levels of interest across the RIASEC scales and this has implications for the determination of the high point codes. This may go part of the way in explaining why some studies of interests within career development theories have produced inconsistent results (see Holland, 1997).

Acknowledgement

I am indebted to the Australian Council for Educational Research through John Ainley, Richard Sweet and Stephen Lamb for making the Youth in Transition data available to me as well as allowing me to work at the ACER. Preparation of this paper was supported financially by the Australian National Training Authority and the Research Centre for Vocational Education and Training.

BEST COPY AVAILABLE

References

Adams, R.J. & Khoo, S.T. (1994). Quest: The interactive test analysis system. Camberwell, Victoria: Australian Council for Educational Research.

Ainley, J., Robinson, L., Harvey-Beavis, A., Elsowrth, G. & Fleming, M. (1994). Subject choice in years 11 and 12. Canberra: Australian Government Publishing Service.

Australian Council for Educational Research (no date). *Overview of the longitudinal surveys of Australian youth program. Technical Paper Number 2*. Melbourne: Australian Council for Educational Research.

Australian Council for Educational Research (no date). *The Youth in Transition 1970 cohort questionnaires: 1980, 1985-1994. Technical Paper Number 5*. Melbourne: Australian Council for Educational Research.

Borgen, F. (1991). Megatrends and milestones in vocational behavior: A 20-year counseling psychology retrospective. *Journal of Vocational Behavior, 39*, 263-290.

Elton, C.F., & Rose, H.A. (1975). A vocational interest test minus sex-bias. *Journal of Vocational Behavior, 7*, 207-214.

Emberton, S.E. (1996). The new rules of measurement. *Psychological Assessment, 8*, 341-349.

Holland, J.L. (1973). *Making vocational choices: A theory of vocational personalities and work environments*. Englewood-Cliffs, NJ: Prentice-Hall.

Holland, J.L. (1985). *Making vocational choices: A theory of vocational personalities and work environments*. (2nd edition). Englewood-Cliffs, NJ: Prentice-Hall.

Holland, J.L. (1997). *Making vocational choices: A theory of vocational personalities and work environments*. (3rd edition). Odessa, Fl.: Psycholgical Assessment Resources.

Lokan, J.J. & Taylor, K.F. (Eds.). (1986). *Holland in Australia: A vocational choice theory in research and practice*. Melbourne: Australian Council for Educational Research.

Reardon, R.C. & Lenz, J.G. (1998). *The Self-Directed Search and Related Holland career materials: A practitioner's guide*. Odessa, Fl.: Psychological Assessment Resources.

Waugh, R.F. (1999). Approaches to studying for students in higher education: A Rasch measurement model analysis. *British Journal of Educational Psychology, 69*, 63-79.

Wright, B. & Masters, G. (1981). *The measurement of knowledge and attitude* (Research Memorandum no. 30). Chicago: Statistical Laboratory, Department of Education, University of Chicago.

Wright, B. & Masters, G. (1982). *Rating scale analysis: Rasch measurement*. Chicago: MESA Press.

BEST COPY AVAILABLE

TABLE 1

Statistics relating to the sub-scales

	Realistic	Investigative	Artistic	Social	Enterprising	Conventional
Mean	+0.00	+0.0	+0.00	+0.00	−0.00	+0.00
SD	+1.06	+0.08	+0.27	+1.12	−0.41	+0.23
Infit ms	+1.01	+0.98	+0.99	+1.00	+1.00	+0.99
Outfit ms	+1.35	+0.98	+0.99	+1.00	+1.00	+1.00
Separability	+0.99	+0.00	+0.93	+0.96	+0.95	+0.85

Means and SD refer to the mean and standard deviation of the scores in logit

TABLE 2

Threshold values (logits) for items on the interest questionnaire

How do you feel about each of these activities?	Difficulty in logits			Infit ms	Outfit ms
	Like very much	Like somewhat	Dislike somewhat or Dislike very much		
Realistic					
1. Working with machines and tools	2.75	0.44	-1.34	0.80	0.82
2. Driving cars	0.43	-2.15	-3.03	1.80	3.12
3. Repairing things	3.07	0.36	-1.78	0.75	0.75
4. Building things	2.82	0.24	-1.81	0.70	0.70
Investigative					
5. Bushwalking	1.80	-0.18	-1.34	1.18	1.20
6. Solving problems and puzzles	1.91	-0.20	-1.72	0.85	0.85
7. Doing all kinds of experiments	1.51	-0.30	-1.53	0.95	0.92
8. Thinking your way through problems	1.88	-0.24	-1.59	0.95	0.94
Artistic					
9. Acting in plays	1.44	0.19	-0.84	0.82	0.81
10. Going to live theatre (e.g. plays)	1.13	-0.10	-1.16	0.80	0.79
11. Doing handcrafts	1.14	-0.44	-1.78	1.44	1.47
12. Writing stories, poems, plays etc	1.33	0.09	-1.00	0.92	0.91
Social					
13. Going shopping	2.65	0.87	-0.53	0.98	0.89
14. Talking with friends	0.65	-2.07	-3.06	1.16	1.04
15. Helping other people	2.02	-0.79	-1.88	0.82	1.05
16. Cooking	2.38	0.43	-0.66	1.03	1.03
Enterprising					
17. Organising things	1.38	-0.87	-2.34	0.98	0.98
18. Selling things to people	2.03	0.16	-1.38	1.16	1.16
19. Managing other people	2.11	0.14	-1.63	0.82	0.82
20. Getting other people to do things your way/influencing others	1.92	0.02	-1.59	1.03	1.03
Conventional					
21. Typing	1.96	0.03	-1.41	1.12	1.14
22. Recording facts and figures	2.03	0.00	-1.84	0.91	0.91
23. Working with figures	1.39	-0.46	-1.94	1.05	1.05
24. Doing office work	1.70	-0.07	-1.34	0.89	0.89

BEST COPY AVAILABLE

FIGURE 1(a)
Realistic Scale

BEST COPY AVAILABLE

FIGURE 1(b)
Investigative Scale

BEST COPY AVAILABLE

FIGURE 1(c)
Artistic Scale

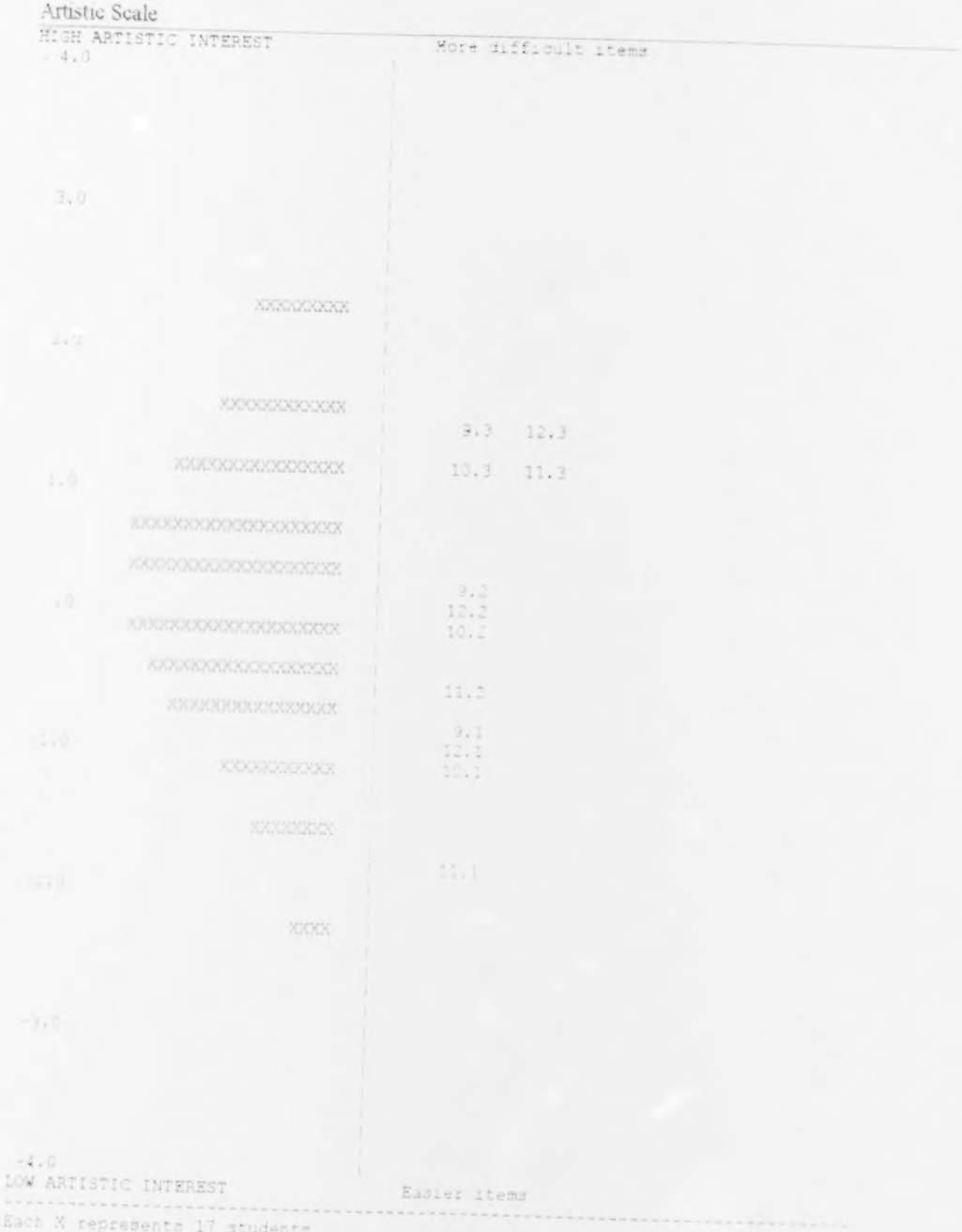


FIGURE 1(d)
Social scale

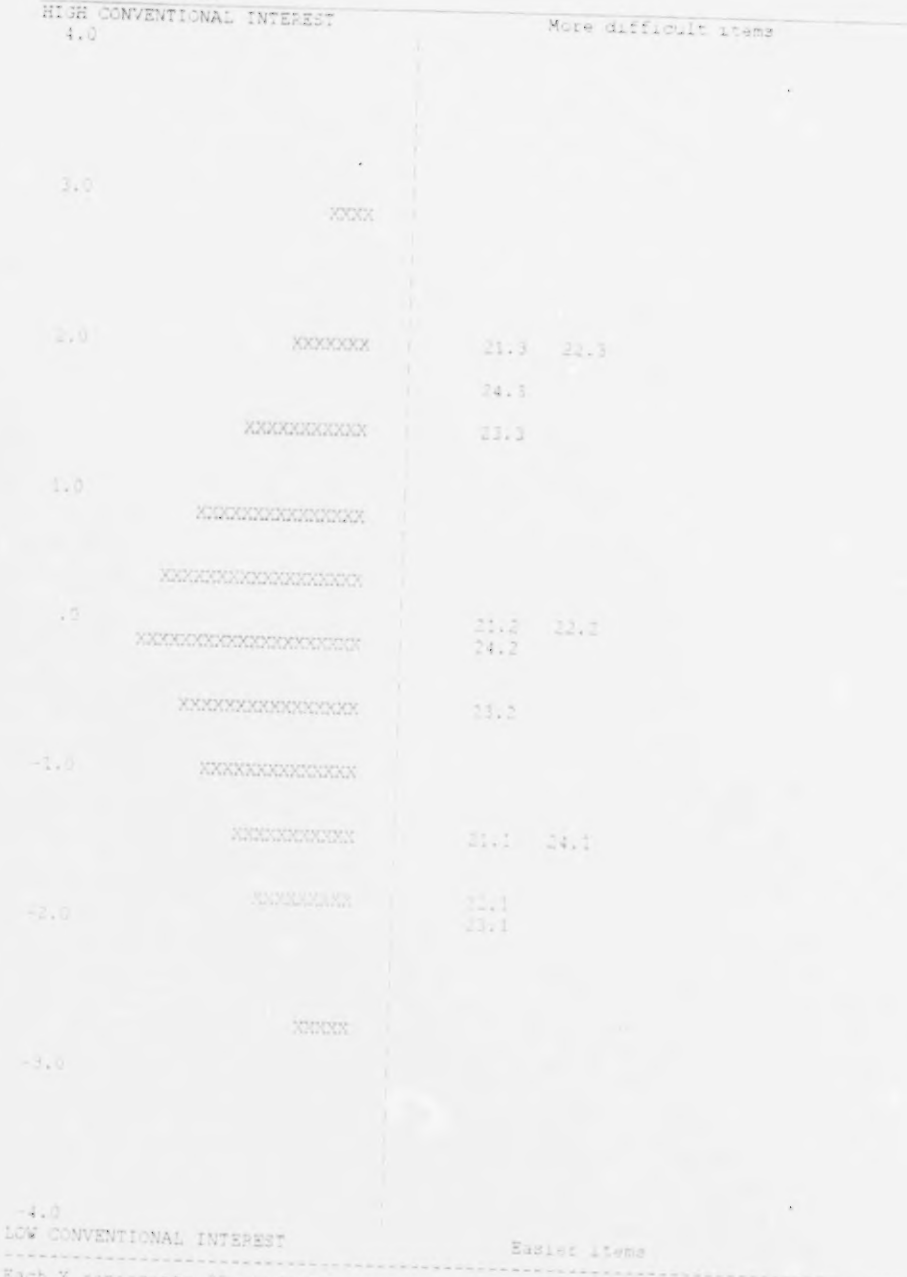
BEST COPY AVAILABLE

FIGURE 1(e)

Enterprising scale

BEST COPY AVAILABLE

FIGURE 1(f)
Conventional scale



Each X represents 20 students

BEST COPY AVAILABLE

Occasional Papers

No. 1. Vocational achievements following spinal cord injury in Australia, August, 1994

No 2. Validity of self-estimates of career interests under matching and non-matching conditions, September 1994

No. 3. Combining school and work - new pathways for learning in Australia, October 1995

No. 4. Examining the validity of Azrin's Job Club methodology as a way of increasing job seeking behaviour and vocational achievements of unemployed persons with spinal cord injury: a case study, February 1996

No. 5. Evaluating the effectiveness of multimedia based learning, July 1998 ED 419 958

No. 6. Repeated judgements of interest in vocational education: A lens model analysis, July 1998 ED 419 959

No. 7. Using a career interest card sort for vocational assessment and counselling, July 1998 ED 419 960

No. 8. The dimensions of vocational interests: A study of Australian high school students, August 1999

No. 9. An evaluation of a model of domain learning: some preliminary data on the effects of students' knowledge, interest and strategies in the acclimation stage of learning, August 1999

No. 10. Analysis of responses to vocational interest items: A study of Australian high school students, August 1999